

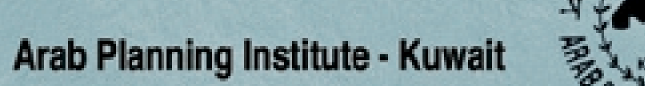
المعهد العربي للتخطيط بالكويت

Arab Planning Institute - Kuwait

منظمة عربية مستقلة

# الإرتباط والانحدار البسيط

سلسلة دورية تعنى بقضايا التنمية في الدول العربية
العدد السابع والأربعون - نوفمبر/تشرين الثاني 2005 - السنة الرابعة






المعهد العربي للتخطيط بالكويت
ص.ب: 5834 الصفاة 13059 - دولة الكويت
هاتف: 4848754 - 4844061 - 4843130 - (965)
فاكس: 4842935

P.O.Box : 5834 Safat 13059 State of Kuwait
Tel : (965) 4843130 - 4844061 - 4848754
Fax : 4842935

E-mail ; api@api.org.kw
web site : http //www.arab-api.org

# أهداف "جسر التنمية"

إن إتاحة أكبر قدر من المعلومات والمعارف لأوسع شريحة من أفراد المجتمع، يعتبر شرطا أساسياً لجعل التنمية قضية وطنية يشارك فيها كافة أفراد وشرائح المجتمع وليس الدولة أو النخبة فقط. وكذلك لجعلها نشاطاً قائماً على المشاركة والشفافية وخاضعاً للتقييم والمساءلة.

وتأتي سلسلة "جسر التنمية " في سياق حرص المعهد العربي للتخطيط بالكويت على توفير مادة مبسطة قدر المستطاع للقضايا المتعلقة بسياسات التنمية ونظرياتها وأدوات تحليلها بما يساعد على توسيع دائرة المشاركين في الحوار الواجب إثارته حول تلك القضايا حيث يرى المعهد أن المشاركة في وضع خطط التنمية وتنفيذها وتقييمها من قبل القطاع الخاص وهيئات المجتمع المدني المختلفة، تلعب دوراً مهماً في بلورة نموذج ومنهج عربي للتنمية يستند إلى خصوصية الأوضاع الاقتصادية والاجتماعية والثقافية والمؤسسية العربية، مع الاستفادة دائماً من التوجهات الدولية وتجارب الآخرين.

*والله الموفق لما فيه التقدم والإزدهار لأمتنا العربية ،،،*

**د. عيسى محمد الغزالي**
**مدير عام المعهد العربي للتخطيط بالكويت**

# قائمة اصدارات "جسر التنمية"

| العنوان | المؤلف | رقم العدد |
|---|---|---|
| مفهوم التنمية | د. محمد عدنان وديع | الأول |
| مؤشرات التنمية | د. محمد عدنان وديع | الثاني |
| السياسات الصناعية | د. أحمد الكواز | الثالث |
| الفقر: مؤشرات القياس والسياسات | د. علي عبدالقادر علي | الرابع |
| الموارد الطبيعية واقتصادات نفاذها | أ. صالح العصفور | الخامس |
| استهداف التضخم والسياسة النقدية | د. ناجي التوني | السادس |
| طرق المعاينة | أ. حسن الحاج | السابع |
| مؤشرات الأرقام القياسية | د. مصطفى بابكر | الثامن |
| تنمية المشاريع الصغيرة | أ. حسان خضر | التاسع |
| جداول المدخلات المخرجات | د. أحمد الكواز | العاشر |
| نظام الحسابات القومية | د. أحمد الكواز | الحادي عشر |
| إدارة المشاريع | أ. جمال حامد | الثاني عشر |
| الاصلاح الضريبي | د. ناجي التوني | الثالث عشر |
| أساليب التنبؤ | أ. جمال حامد | الرابع عشر |
| الادوات المالية | د. رياض دهال | الخامس عشر |
| مؤشرات سوق العمل | أ. حسن الحاج | السادس عشر |
| الاصلاح المصرفي | د. ناجي التوني | السابع عشر |
| خصخصة البنى التحتية | أ. حسان خضر | الثامن عشر |
| الأرقام القياسية | أ. صالح العصفور | التاسع عشر |
| التحليل الكمي | أ. جمال حامد | العشرون |
| السياسات الزراعية | أ. صالح العصفور | الواحد والعشرون |
| اقتصاديات الصحة | د. علي عبدالقادر علي | الثاني والعشرون |
| سياسات أسعار الصرف | د. بلقاسم العباس | الثالث والعشرون |
| القدرة التنافسية وقياسها | د. محمد عدنان وديع | الرابع والعشرون |
| السياسات البيئية | د. مصطفى بابكر | الخامس والعشرون |
| إقتصاديات البيئة | أ. حسن الحاج | السادس والعشرون |
| تحليل الأسواق المالية | أ. حسان خضر | السابع والعشرون |
| سياسات التنظيم والمنافسة | د. مصطفى بابكر | الثامن والعشرون |
| الأزمات المالية | د. ناجي التوني | التاسع والعشرون |
| إدارة الديون الخارجية | د. بلقاسم العباس | الثلاثون |
| التصحيح الهيكلي | د. بلقاسم العباس | الواحد والثلاثون |
| نظم البناء والتشغيل والتحويل B.O.T. | د. أمل البشبيشي | الثاني والثلاثون |
| الاستثمار الأجنبي المباشر : تعاريف | أ. حسان خضر | الثالث والثلاثون |
| محددات الاستثمار الأجنبي المباشر | د. علي عبدالقادر علي | الرابع والثلاثون |
| نمذجة التوازن العام | د. مصطفى بابكر | الخامس والثلاثون |
| النظام الجديد للتجارة العالمية | د. أحمد الكواز | السادس والثلاثون |
| منظمة التجارة العالمية: إنشاؤها وآلية عملها | د. عادل محمد خليل | السابع والثلاثون |
| منظمة التجارة العالمية: أهم الإتفاقيات | د. عادل محمد خليل | الثامن والثلاثون |
| منظمة التجارة العالمية: آفاق المستقبل | د. عادل محمد خليل | التاسع والثلاثون |
| النمذجة الإقتصادية الكلية | د. بلقاسم العباس | الأربعون |
| تقييم المشروعات الصناعية | د. أحمد الكواز | الواحد والأربعون |
| المؤسسات والتنمية | د. عماد الامام | الثاني والأربعون |
| التقييم البيئي للمشاريع | أ. صالح العصفور | الثالث والأربعون |
| مؤشرات الجدارة الإئتمانية | د. ناجي التوني | الرابع والأربعون |
| الدمج المصرفي | أ. حسان خضر | الخامس والأربعون |
| اتخاذ القرارات | أ. جمال حامد | السادس والأربعون |
| **الإرتباط والانحدار البسيط** | **أ. صالح العصفور** | **السابع والأربعون** |

للاطلاع على الأعداد السابقة يمكنكم الرجوع إلى العنوان الإلكتروني التالي :

http://www.arab-api.org/develop_1.htm

# المحتويـــات

# الارتباط والانحدار البسيط

إعداد: أ. صالح العصفور

## أولا - الانحدار الخطي والارتباط

يكتسب قياس وتحديد درجة واتجاه العلاقة بين المتغيرات اهمية كبيرة في فهم الظواهر بمختلف أنواعها. وعندما يكون الأمر متعلقاً بمتغير واحد، فإن مقاييس النزعة المركزية تصف لنا القيمة التي تقع في مركز مجموعة البيانات، كما تصف لنا مقاييس التشتت درجة انتشار وتبعثر وتوزيع قيم هذه البيانات. ولكن عندما يتعلق الأمر بمتغيرين أو أكثر، فان الباحث يتطلع إلى قياس وتحديد درجة واتجاه العلاقة بين المتغيرات أو الظواهر قيد الدراسة. ويقوم بعد ذلك باستخدام العلاقات الموجودة في التنبؤ بقيمة أحد المتغيرات بدلالة المتغير (المتغيرات) الأخرى. ومحاولة التعبير عن هذه العلاقات بدالة خطية، لأن وجود علاقة أو ارتباط بين المتغيرات لا يعني بالضرورة إمكانية التعبير عن هذه العلاقة بشكل خطي. ويعبر عن هذه العلاقة بالأرقام أو القيم الكمية كما قد يعبر عنها بالوصف. فالإحصائي عندما يتحدث عن إحدى العلاقات الدالية بين متغيرين فانه يقصد بذلك أن المتغيرين يمكن ربطهما بمعادلة رياضية، بحيث إذا علمت قيمة أحدهما (المتغير المستقل) أمكن معرفة قيمة المتغير الآخر(المتغير التابع) .

**يُعنى الانحدار البسيط بدراسة العلاقة بين متغيرين على هيئة علاقة دالية، بحيث يمكن الاعتماد على المعلومات المتوفرة عن أحدهما للتنبؤ عن الآخر.**

لنضرب مثلاً المكالمات الهاتفية ، فإذا عرفنا أن قيمة المكالمة الواحدة هي 20 فلساً، وأن استخدام الهاتف كان لخمسة مرات، فان القيمة التي يجب دفعها هي 100 فلس. ويمكن التعبير عن ذلك بطريقة رياضية، حيث نرمز لقيمة المكالمات بالرمز Y ولعدد مرات استعمال الهاتف بالرمز X ، وبذلك تكون العلاقة الرياضية :

$$Y = 20\ X$$

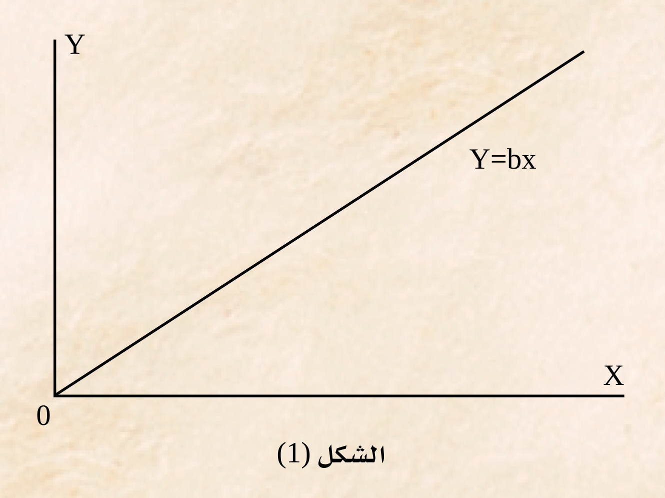
وتبقى هذه الصيغة صالحة مادامت قيمة المكالمة الواحدة ثابتة أي مساوية لعشرين فلساً. واذا ما زادت قيمة المكالمة إلى 25 فلساً فان المعادلة تصبح :

$$Y = 25\ X$$

وعليه فإنه يمكن تعميم القاعدة على كل القيم التي تبلغها المكالمة الهاتفية (b) فتصبح الصياغة الرياضية في هذه الحالة كما في المعادلة :

$$Y = b\ X$$

ولزيادة الإيضاح، يمكن التعبير عن هذه المعادلة بالرسم البياني، حيث يخصص محور الإحداثيات الأفقية لقيم X ومحور الإحداثيات الرأسية لما يقابلها من قيم Y وبذلك نحصل على الشكل (1)، فتبقى b ثابتة في المعادلة يمثلها الخط المعروف بخط الانحدار، وفيه نلاحظ أن كل قيمة من قيم X تقابلها قيمة من قيم Y وان قيمة Y تساوي صفراً عندما تكون X مساوية للصفر. وينطلق هذا الخط الانحداري من أصل الإحداثيات الرأسية والأفقية أي من نقطة تقاطعهما وهي نقطة الصفر .

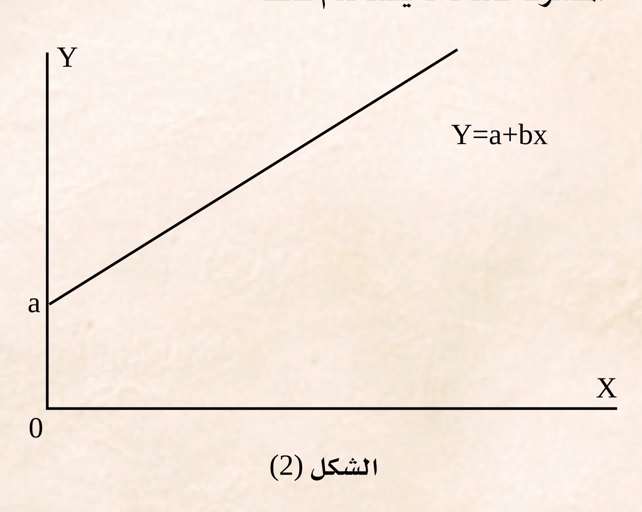



الشكل (1)

ولكن صاحب الهاتف لا يدفع لشركة الاتصالات قيمة المكالمات الهاتفية فقط، بل يقوم بدفع قيمة الاشتراك التي تضاف إلى قيمة المكالمات الهاتفية. وهذا الاشتراك يدفعه المشترك سواء استعمل هاتفه أم لم يستعمله. وعليه فإننا يجب أن نضيف إلى معادلتنا السابقة الرمز (a) للدلالة على قيمة الاشتراك الثابتة، فتصبح المعادلة التالية كما يلي :

$$Y = a + bX$$

ويعبر عن هذه المعادلة بالرسم البياني المبين في {الشكل (2)} حيث يلاحظ ان الخط الانحداري لا يمر بأصل الإحداثيات كما في الشكل السابق بل يقطع مستقيم الإحداثيات الرأسية في نقطة تساوي قيمة الاشتراك. أي أن نقطة (a) تساوي القيمة التي يدفعها المشترك عندما لا يستخدم هاتفه .



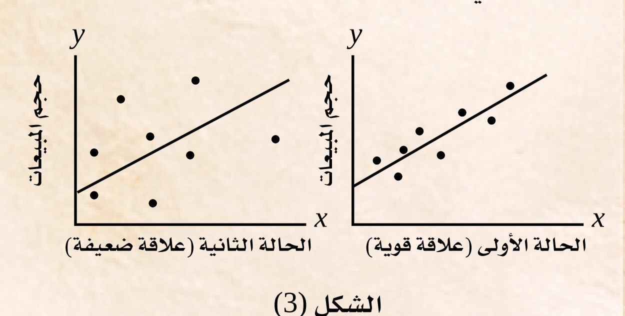

الشكل (2)

لكن العلاقة الدالية لا يشترط فيها دائماً أن تتبع خطاً مستقيماً بل قد تتبع خطاً منحنياً. أي أنها لا تكون دائماً بسيطة بل قد تكون مركبة ، ولحساب قيمة العلاقة بين المتغيرين X و Y فيجب معرفة قيمة المعاملين a و b ، وبمعرفة قيمة المتغير المستقل يستخرج قيمة المتغير التابع Y أو العكس .

ولكن ما يحصل في الواقع قد لا يكون بالشكل الذي افترضناه في مثالنا بخصوص الهاتف . فالعلاقة بين الظاهرتين قد لا تكون خطية بشكل كامل . ولبيان ذلك يفترض أن مشكلتنا هي في بيان أو إيجاد العلاقة بين حجم المبيعات وسنوات خبرة الباعة (مندوبي المبيعات) . فلو قام المدير بتسجيل بيانات تتضمن حجم المبيعات السنوية وسنوات الخبرة لمندوبي المبيعات . وكانت هذه البيانات كما في الجدول (1):

| مندوبي المبيعات | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| سنوات الخبرة | 1 | 3 | 4 | 4 | 6 | 8 | 10 | 10 | 11 | 13 |
| حجم المبيعات (ألف دينار) | 80 | 97 | 92 | 102 | 103 | 111 | 119 | 123 | 117 | 136 |

إن الخطوة الأولى للبحث عن علاقة هي رسم بياني للبيانات أعلاه على الشكل أدناه (3) حيث مثلت سنوات الخبرة على الإحداثي الأفقي والمبيعات السنوية على الإحداثي الرأسي. ونكون بذلك قد حصلنا على رسم انتشاري . وقد أعطي هذا الاسم نظراً لانتشار وتبعثر النقاط على الشكل أو الرسم . وهذا الرسم الانتشاري (Scatter Diagram) المبين أدناه يساعد على تكوين استنتاج مبدئي حول إمكانية وجود علاقة بين المتغيرات، حيث يبين هذا الرسم في دراستين مختلفتين وجود علاقة قوية في الحالة الأولى، وعلاقة ضعيفة في الحالة الثانية.



الشكل (3)

وعموماً لجأ الإحصائيون إلى تصنيف المتغيرات إلى متغيرات مستقلة ومتغيرات تابعة. ويستخدم هذا التصنيف من أجل تحديد المتغير المفسّر (المستقل) والمتغير المفسَّر (التابع). وفي مثالنا هذا فانه يشار إلى سنوات الخبرة كمتغير مستقل، ويستخدم للتنبؤ بحجم المبيعات أو المتغير التابع. وتجدر الملاحظة هنا أنه جرت العادة في الرسم الانتشاري أن يكون المتغير المستقل على الإحداثي الأفقي والمتغير التابع على الإحداثي الرأسي.

ولو نظرنا إلى الشكل أعلاه فهو يعطينا لمحة عن البيانات اذ يشير إلى أن هناك فرصة لوجود علاقة بين المتغيرات. حيث أن سنوات الخبرة المنخفضة تترافق مع انخفاض حجم المبيعات، و يترافق ارتفاع سنوات الخبرة في الغالب مع ارتفاع حجم المبيعات السنوية. وتبدو هذه العلاقة بين المتغيرين خطاً مستقيماً تقريباً أو معادلة خطية. وسنبين لاحقاً كيفية تطوير مثل هذه العلاقة الخطية باستخدام الأسلوب المسمى طريقة المربعات الصغرى .

## ثانيا . خط الانحدار البسيط:

سنحاول هنا التركيز على مهمة إيجاد خط مستقيم يمثل الرسم الانتشاري للبيانات أفضل تمثيل. أي أننا سنوفق معادلة على الشكل التالي:

$$\hat{y} = a + bx$$

حيث:

$\hat{y}$ = القيمة المقدرة للمتغير التابع.

$x$ = قيمة المتغير المستقل.

$a$ = تقاطع المحور y (أي أنها قيمة عندما تكون x=0).

$b$ = تغير المتغير التابع yنتيجة لتغير المتغير المستقل (ميل خط الانحدار).

**إذا لم تكن العلاقة بين المتغيرين علاقة دالية بل علاقة ترابط فقط (أي أن التغير في إحداهما لا يسبب التغير في الآخر) فإن مهمة الإحصاء هي قياس درجة العلاقة بين المتغيرين واتجاهها.**

في تحليل الانحدار تكون المعادلة التي نستخدمها لوصف البيانات هي معادلة الانحدار المقدرة. وتجدر الإشارة هنا إلى أننا سنركز في هذا الجزء على معادلات خط الانحدار التي تأخذ شكل خط مستقيم. وعليه فانه يشار إلى معادلة خط الانحدار بأنها خط الانحدار المقدر. ويشار عموماً

إلى علاقة الخط المستقيم المعنية بمتغيرين أحدهما مستقل والآخر تابع على أنها انحدار خطي بسيط.

ولتوضيح فكرة خط الانحدار البسيط، نأخذ المثال المقدم في جدول (1) بخصوص حجم المبيعات السنوية وسنوات الخبرة لعشرة مندوبي مبيعات. فباستخدام الرسم الانتشاري المبين في شكل (3)، كيف يمكن اختيار أفضل خط لوصف العلاقة بين حجم المبيعات وسنوات الخبرة؟

يتم في البداية وضع خط يعتقد أنه الأقرب لإحداثيات البيانات المرسومة بحيث نتمكن من تحديد بداية الخط (a). إضافة إلى ذلك، فإننا باستعمال أي نقطتين، يمكننا تحديد الميل (b). وإذا ما كلفنا شخصين (باحثين) ليقوما بهذه المهمة، فانهما سيخرجان بمعادلتي خط انحدار مختلفتين ، ولتكونا كما يلي:

الباحث الأول : $\hat{y} = 78 + 4.2x$

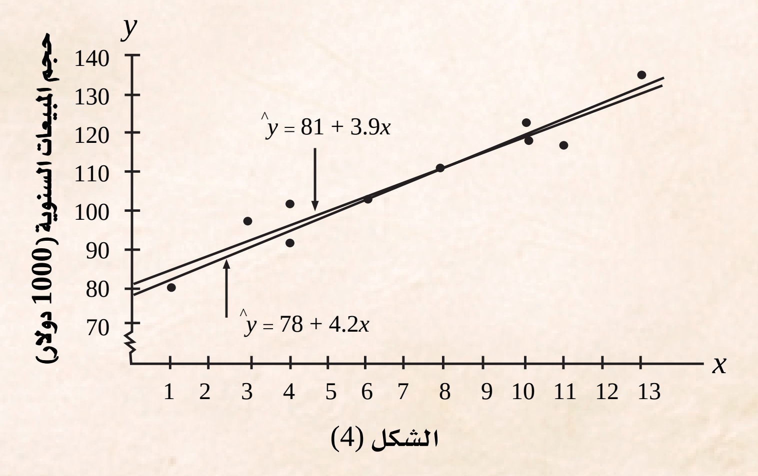
الباحث الثاني : $\hat{y} = 81 + 3.9x$

حيث:

$\hat{y}$ = القيمة المقدرة للمبيعات السنوية (بآلاف الدنانير) .

x = سنوات الخبرة في مجال المبيعات .

يبين الشكل (4) مدى ملاءمة هذين الخطين المستقيمين للبيانات الموضحة في الشكل (3). ولكن رغم أن الخطين مشابهين إلى حد كبير للخط الذي يمكن أن نرسمه ، إلا أننا لا نمتلك معياراً لاختيار أفضل خط. إن اختلاف نظرة الأشخاص لنفس البيانات وخروجهم بعلاقات رياضية مختلفة يمكن أن يؤثر على صانعي القرار. فلو وضع شخص آخر خطاً مختلفاً، ربما يتساءل صانع القرار ، هل يمكن أن يكون هناك قرار آخر؟ وعليه فإننا بحاجة إلى اتفاق على معيار أو مقياس معين لاختيار خط الانحدار المقدر بحيث أن أي شخص آخر يقوم بتحليل البيانات سوف يختار نفس هذا الخط.



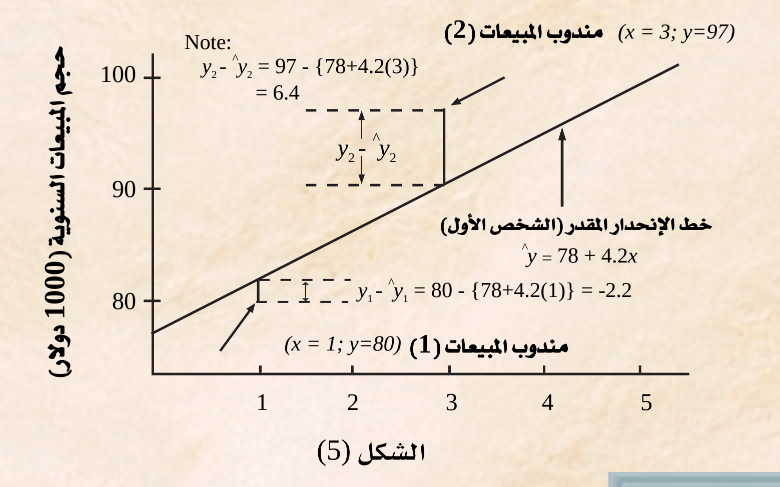

الشكل (4)

## ثالثا ـ طريقة المربعات الصغرى :

تعطي هذه الطريقة أكفأ تقدير لمعادلة الانحدار الخطي، بحيث تجعل من مجموع مربعات الاختلافات بين القيم المشاهدة للمتغير التابع (Y) والقيم المقدرة لهذا المتغير ($\hat{y}$) عند نهايتها الصغرى. ولنلق نظرة متفحصة على مدلول مجموع مربعات الاختلافات بين القيم المشاهدة والقيم المقدرة للمتغير التابع ، ومعناها الحقيقي.

ومن أجل مزيد من التوضيح سنسلط الضوء على جزء من الرسم الانتشاري فقط لأول مشاهدتين والقيم المقدرة لهما على خط الانحدار $\hat{y} = 78 + 4.2x$ وبالنظر إلى الشكل التالي (5)، يمكن ملاحظة الفرق بين القيمة المشاهدة من المبيعات السنوية لمندوب المبيعات الأول ($y_1$) والقيمة المقدرة لهذه المبيعات ($\hat{y}_1$) هو كما يلي :

$$y_1 - \hat{y}_1 = 80 - 82.2 = -2.2$$




الشكل (5)

جدول (2): حساب مجموع مربعات الفروق (الاختلافات) لخط الانحدار المقدر $\hat{y} = 78 + 4.2x$

| مندوب المبيعات | سنوات الخبرة $(X_i)$ | القيم المشاهدة لحجم المبيعات $(y_i)$ | القيم المقدرة للمبيعات $(\hat{y} = 78 + 4.2x)$ | الفروق بين القيم المشاهدة والمقدرة $(y_i - \hat{y}_i)$ | مربعات الفروق $(y_i - \hat{y}_i)^2$ |
|---|---|---|---|---|---|
| 1 | 1 | 80 | 78+4.2(1)=82.2 | -2.2 | 4.84 |
| 2 | 3 | 97 | 78+4.2(3)=90.6 | 6.4 | 40.96 |
| 3 | 4 | 92 | 78+4.2(4)=94.8 | -2.8 | 7.84 |
| 4 | 4 | 102 | 78+4.2(4)=94.8 | 7.2 | 51.84 |
| 5 | 6 | 103 | 78+4.2(6)=103.2 | -0.2 | 0.04 |
| 6 | 8 | 111 | 78+4.2(8)=111.6 | -0.6 | 0.36 |
| 7 | 10 | 119 | 78+4.2(10)=120.0 | -1.0 | 1.00 |
| 8 | 10 | 123 | 78+4.2(10)=120.0 | 3.0 | 9.00 |
| 9 | 11 | 117 | 78+4.2(11)=124.2 | -7.2 | 51.84 |
| 10 | 13 | 136 | 78+4.2(13)=132.6 | 3.4 | 11.56 |
| المجموع | | | | | 179.28 |

وبالتالي فان مربع الفرق سيكون :

$(y_1 - \hat{y}_1)^2 = (-2.2)^2 = 4.84$

وسيكون الفرق للمشاهدة الثانية $(y_2 - \hat{y}_2)^2 = 6.4$

ومربع الفرق $(y_2 - \hat{y}_2)^2 = 40.96$

ويبين الجدول التالي (2) بقية الحسابات للمشاهدات الثمانية الأخرى .

من جدول (2) نلاحظ أن مجموع مربعات الاختلافات = 179.28. وقد قمنا أيضاً باحتساب مجموع مربعات الاختلافات للخط المقدر الآخر: $\hat{y} = 81 + 3.9x$ والذي أعد من قبل الباحث الثاني في مثالنا أعلاه. وقد كانت النتيجة 172.32. وحيث أن مجموع المربعات لهذا الخط المقدر هي أصغر أو أقل من الخط الأول فإننا نعتبر هذا الخط أكفأ لتقدير خط الانحدار من الخط الأول . ولكن طريقة المربعات الصغرى لا تعتبر هذين الخطين هما الخطين الوحيدين. فهي التي توجد في الحقيقة خط الانحدار المقدر الذي يعطي أصغر مجموع لمربعات الفروق (الاختلافات) لكل الخطوط المختارة .

**1- تقدير خط الانحدار بطريقة المربعات الصغرى :**

عند استعمال طريقة المربعات الصغرى لتقدير معالم خط الانحدار المقدر ، فان الإحصائيين يرون بأن أفضل قيم لكل من a (ثابت الانحدار) و b (معامل الانحدار) يمكن إيجادها باستعمال المعادلات التالية :

$$b = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \quad \text{.............(1)}$$

$$a = \bar{y} - b_1 \bar{x} \quad \text{.............(2)}$$

حيث أن :

$x_i$ = قيم المتغير المستقل لـ (i) من المشاهدات .

$y_i$ = قيم المتغير التابع لـ (i) من المشاهدات .

$\bar{x}$ = قيمة متوسط المتغير المستقل .

$\bar{y}$ = قيمة متوسط المتغير التابع .

$n$ = عدد المشاهدات .

ويبين الجدول رقم (3) بعض الحسابات الضرورية لاحتساب خط الانحدار المقدر في المثال السابق.

جدول (3): بعض الحسابات الضرورية
لتقدير معاملات خط الانحدار

| مندوب المبيعات (i) | سنوات الخبرة ($x_i$) | $y_i$ | $x_i y_i$ | $x^2_i$ |
|---|---|---|---|---|
| 1 | 1 | 80 | 80 | 1 |
| 2 | 3 | 97 | 291 | 9 |
| 3 | 4 | 92 | 368 | 16 |
| 4 | 4 | 102 | 408 | 16 |
| 5 | 6 | 103 | 618 | 36 |
| 6 | 8 | 111 | 888 | 64 |
| 7 | 10 | 119 | 1190 | 100 |
| 8 | 10 | 123 | 1230 | 100 |
| 9 | 11 | 117 | 1287 | 121 |
| 10 | 13 | 136 | 1768 | 169 |
| المجموع | 70 | 080,1 | 128,8 | 632 |

وباستعمال الأرقام والقيم الموجودة في الجدول (3) وباستخدام الصيغ الرياضية 1 و 2 الآنفة الذكر ، يمكننا احتساب الميل (b) ونقطة تقاطع خط الانحدار مع (y) وهي a كما يلي :

(1) احتساب $\bar{x}$ و $\bar{y}$ :

$$\bar{x} = \frac{\sum x_i}{n} = \frac{70}{10} = 7$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{1080}{10} = 108$$

(2) احتساب الميل (b):

$$b = \frac{n\sum x_i \sum y_i - \sum x_i y_i}{n\sum x^2_i - (\sum x_i)^2}$$

$$= \frac{10(8,128) - 70(1080)}{10(632) - (70)^2}$$

$$= \frac{5680}{1420} = 4$$

(3) احتساب تقاطع (a):

$$a = \bar{y} - b_1 \bar{x}$$

$$= 108 - 4(7)$$

$$= 80$$

وعليه تكون معادلة خط الانحدار المقدرة كما يلي :

$$\hat{y} = 80 + 4x$$

بملاحظة أن الميل b موجب ، هذا يعني أنه كلما زادت سنوات الخبرة (x) لدى العاملين في المبيعات كلما زاد حجم المبيعات (y). ويتضح من المسألة الموجودة بين أيدينا أن هناك علاقة إيجابية بين x و y . ولكن في حالات أخرى قد يكون الميل سالباً ليشير إلى أنه كلما زادت x انخفضت y ، وفي هذه الحالة تكون العلاقة بين x و y سالبة أو عكسية .

**في حالة العلاقة الدالية فإن علاقة الارتباط حتمية، ولكن وجود الارتباط بين متغيرين لا يعني بالضرورة وجود علاقة دالية.**

ولو قمنا باحتساب مجموع مربعات الفروق (الاختلافات) بين القيم المشاهدة والقيم المقدرة استناداً إلى خط الانحدار المقدر بطريقة المربعات الصغرى 80 + $\hat{y} = 4x$، فإننا نجد أن $\sum(y - \hat{y}_1)^2 = 170$ وهي قيمة أقل من مجموع المربعات للخطين اللذين نوقشا سابقاً. ومن المهم الإشارة هنا إلى أن طريقة المربعات الصغرى تضمن

عدم وجود أي خط آخر يمكن أن يؤمن مجموع مربعات أصغر من 170 في مثالنا الحالي . لذلك نعتبر أن طريقة المربعات الصغرى هي التي تضمن أكفأ تقدير لخط الانحدار.

### 2- استخدام معادلة الانحدار الخطي البسيط في التنبؤ:

مع الاعتقاد بأن معادلة خط الانحدار المقدرة بطريقة المربعات الصغرى تصف العلاقة بين x و y بشكل دقيق ، فانه يبدو أنه من المعقول استخدام هذه المعادلة الرياضية في تقدير قيم y عند معرفة قيم x .وفي هذه الحالة فإن معادلة خط الانحدار المقدرة تسمى معادلة خط انحدار y على x. أما إذا كان المتغير التابع هو x والمتغير المستقل هو y فإن المعادلة المقدرة ستكون معادلة انحدار x على y، وبالتالي فإننا سنتنبأ من خلالها بقيم x بعد معرفة قيم y .

في مثالنا الحالي المتعلق بمسألة حجم المبيعات السنوية وسنوات الخبرة ، فانه يمكننا استخدام معادلة خط الانحدار المقدرة بطريقة المربعات الصغرى $\hat{y} = 80 + 4x$ في تقدير حجم المبيعات السنوية المتوقعة لثلاثة متقدمين لوظيفة مندوب مبيعات . ويبين الجدول (4) نتائج عملية التقدير .

من الجدول السابق نرى أن تحليل الانحدار السابق يعطي دعماً لقرار أن السيد رشوان محمد هو الشخص الأفضل كمندوب مبيعات ، حيث يمكن تقدير مبيعاته خلال السنة الأولى من التحاقه بالشركة بحوالي 100,000 دولار ، ويفوق بذلك المبيعات المقدره لعلي مرتجى المتقدم الثاني للوظيفة بحوالي 12,000 دولار. في هذا المثال لمسنا أحد أهم استخدامات تحليل الانحدار : بتزويد صيغة رياضية تدعم بالمعلومات صانعي القرار في المجال التجاري والاقتصادي .

**هناك الكثير من العلاقات الزائفة بين بعض المتغيرات، فبالرغم من وجود علاقة ارتباط قوية، إلاّ أن هذه العلاقة لا تعني بالضرورة وجود متغير تابع ومتغير مستقل.**

### 3- جودة التقدير :

الخطوة التالية في دراسة الانحدار هي قياس جودة التقدير في نموذج خط الانحدار . ولكن قبل ذلك لا بد من وضع مجموعة من الفرضيات (في قيم البيانات) تمكننا من إنجاز تحليلات إحصائية إضافية .

جدول (4): استخدام معادلة خط الانحدار في تقدير حجم المبيعات السنوية لثلاثة متقدمين

| المتقدم للوظيفة | سنوات الخبرة | حجم المبيعات السنوية المقدرة (بالألف دولار) |
|---|---|---|
| رشوان محمد | 5 | 100 = (5) 4 + 80 |
| علي مرتجى | 2 | 88 = (2) 4 +80 |
| يحيى الربيعي | 0 | 80 = (0) 4 + 80 |

يفترض النموذج أن قيم المتغير المستقل x يمكن أن تقابلها قيم للمتغير التابع y . وبالتالي فان تغير y ناتج عن (1) التغير في x (2) التغير الباقي e وهو تغير عشوائي غير مفسر من قبل النموذج.

سوف نستعرض هنا في هذا الجزء طريقة إحصائية لقياس أو وصف جودة خط الانحدار المقدر . وقد كنا قد استعرضنا طريقة المربعات الصغرى كتقنية من أجل تقليل مجموع مربعات الفروق (الاختلافات) بين القيم المشاهدة للمتغير التابع ($y_i$) والقيم المقدرة له ($\hat{y}_i$). ويلاحظ أن الفروق تمثل في الحقيقة خطأ استعمال $\hat{y}_i$ كقيم مقدرة لـ $y_i$. وعليه فان مجموع المربعات الناتجة يمكن أن يشار إليها على أنها مجموع مربعات الأخطاء ، وسوف نرمز إليها بـ SSE.

حيث أن : $SSE = \sum(y_i - \hat{y}_i)^2$

وبالعودة إلى مثالنا السابق المتعلق بحجم المبيعات وسنوات الخبرة في البيع فقد كانت $\sum(y_i - \hat{y}_1)^2 = 170$ . وعليه فان SSE = 170 هي عبارة عن مقدار الخطأ الناتج عن استخدام خط الانحدار المقدر ($\hat{y} = 80 + 4x$) لتقدير قيم y.

**تسمى النسبة بين الاختلافات المفسَّرة والاختلاف الكلي بمعامل التحديد، فإذا كانت تساوي صفراً فإن ذلك يعني أن الاختلاف الكلي جميعه غير مفسَّر.**

ولكننا لو سئلنا عن تقدير حجم المبيعات السنوية لمندوب المبيعات دون اللجوء إلى أو دون معرفة سنوات الخبرة لكل مندوب وبدون اللجوء إلى معادلة خط الانحدار ، فان متوسط مبيعات العينة Y وهو $Y = 108$، سيكون هو التقدير الأمثل بالنسبة لنا . والآن انظر إلى الأخطاء أو الفروق التي حصلنا عليها لو استعملنا المتوسط لكل مندوبي المبيعات (انظر الجدول 5). إن مجموع مربعات التشتت للمشاهدات الحقيقية حول المتوسط (108) هو 2,442. وهذه القيمة تمثل مجموع المربعات الكلي للفروق قبل تحليل الانحدار ، ويشار إليها عموماً على أنها المجموع الكلي لمربعات الاختلاف حول المتوسط . وسوف نرمز لها بـ SST . وفي مثالنا السابق سيكون المجموع الكلي لمربعات الاختلاف :

$$SST = \sum(Y_i - \bar{Y})^2 = 2442$$

جدول (5): مجموع مربعات الاختلاف حول المتوسط (108) SST

| مندوب المبيعات | سنوات الخبرة ($x_i$) | ($y_i$) | ($y_i - \bar{y}$) | $(y_i - \bar{y})^2$ |
|---|---|---|---|---|
| 1 | 1 | 80 | -28 | 784 |
| 2 | 3 | 97 | -11 | 121 |
| 3 | 4 | 92 | -16 | 256 |
| 4 | 4 | 102 | -6 | 36 |
| 5 | 6 | 103 | -5 | 25 |
| 6 | 8 | 111 | 3 | 9 |
| 7 | 10 | 119 | 11 | 121 |
| 8 | 10 | 123 | 15 | 225 |
| 9 | 11 | 117 | 9 | 81 |
| 10 | 13 | 136 | 28 | 784 |
| $\sum$ | | | | 2242 |

مجموع مربعات التشتت للمشاهدات الحقيقية حول المتوسط (108) هو 2,442. وهذه القيمة تمثل مجموع المربعات الكلي للفروق قبل تحليل الانحدار ، ويشار إليها عموماً على أنها المجموع الكلي لمربعات الاختلاف حول المتوسط . وسوف نرمز لها بـ SST . وفي مثالنا السابق سيكون المجموع الكلي لمربعات الاختلاف :

$$SST = \sum(Y_i - Y)^2 = 2442$$

فلو كان مقدار الاختلاف هو 2442 (SST) قبل تحليل الانحدار و SSE (170) بعد تحليل الانحدار، فيمكن الاستنتاج أن الفرق (2272) هو مجموع مربعات

الاختلاف (التباين) المفسرة بواسطة معادلة خط الانحدار. وعموماً يدعى مجموع المربعات هذا بمجموع مربعات الانحدار ويرمز له بـ SSR وهو الاختلاف المفسر. ورغم أن الطريقة أعلاه هي الطريقة المستخدمة عادة في احتساب الاختلاف المفسر (SSR)، إلا أنه يمكن تبيان طريقة مباشرة لاحتساب SSR باستخدام الصيغة التالية :

$$SSR = \sum(\hat{Y}_i - Y)^2$$

واحتساب (SSR) باستخدام الصيغة أعلاه مبين في جدول (6)، وقد تم الحصول على قيمة (2272=SSR) وهي قيمة مطابقة لما تم الحصول عليه أعلاه. والعلاقة بين SSE، SST و SSR تشكل أساساً لواحدة من أهم النظريات في الإحصاء التطبيقي. وتقول هذه النظرية، بشكل عام : أن مجموع مربعات المشاهدات حول متوسطها (SST) يمكن تجزئتها إلى جزئين : SSE و SSR بحيث : SST = SSE + SSR

والآن، لنرى كيف يمكن استعمال العلاقة المبينة أعلاه في تطوير مقياس لجودة التوفيق لمعادلة خط الانحدار المقدرة.

**يكون التوفيق كاملاً لمعادلة خط الانحدار إذا كانت جميع المشاهدات واقعة على خط مستقيم، وعندها يمر خط الانحدار المقدر بطريقة المربعات الصغرى في جميع النقاط.**

سيكون لدينا توفيق كامل لخط الانحدار المقدر لو كانت كل المشاهدات تقع على خط مستقيم. وفي هذه الحالة فان خط الانحدار المقدر بطريقة المربعات الصغرى سيمر في جميع النقاط. ولذلك ستكون SSE=0. وسيكون SSR = SST، وبالتالي SST/SSR = 1. ومن جانب آخر فان توفيقاً ضعيفاً للبيانات المشاهدة ينتج عن SSE كبيرة. وحيث أن SST = SSE + SSR، فان أكبر

جدول (6): الاحتساب المباشر لمجموع مربعات الانحدار (SSR)

| مندوب المبيعات | سنوات الخبرة $x_i$ | حجم المبيعات $y_i$ | حجم المبيعات المقدرة $(\hat{y} = 80 + 4x_i)$ | $\hat{y}_i - \bar{y}$ | $(\hat{y}_i - \bar{y})^2$ |
|---|---|---|---|---|---|
| 1 | 1 | 80 | 84 | -24 | 576 |
| 2 | 3 | 97 | 92 | -16 | 256 |
| 3 | 4 | 92 | 96 | -12 | 144 |
| 4 | 4 | 102 | 96 | -12 | 144 |
| 5 | 6 | 103 | 104 | -4 | 16 |
| 6 | 8 | 111 | 112 | 4 | 16 |
| 7 | 10 | 119 | 120 | 12 | 144 |
| 8 | 10 | 123 | 120 | 12 | 144 |
| 9 | 11 | 117 | 124 | 16 | 256 |
| 10 | 13 | 136 | 132 | 24 | 576 |
| Σ | | | | | 2272 |

SSE (وبالتالي أضعف جودة توفيق) سيحدث عندما تكون SST = SSE ، وفي هذه الحالة فان SSR = صفر ، وعندها لن يكون لخط الانحدار المقدر أي دور في المساعدة بتخمين قيم y. وهكذا فان أسوأ توفيق ممكن لخط الانحدار يكون عندما تكون SSR = صفر وعندها تكون نسبة SST/SSR = صفر .

وإذا ما استخدمنا نسبة SST/SSR لتقييم مدى جودة التقدير لخط الانحدار ، سيكون لدينا مقياس يمكن أن يأخذ قيماً بين صفر و 1، وكلما اقتربت القيمة من 1 يدل على جودة أفضل في التقدير لخط الانحدار. والكسر الناتج عن SST/SSR يسمى معامل التحديد ويرمز له بالرمز 2R. وهو عبارة عن النسبة بين الاختلافات المفسرة والاختلاف الكلي .

$$\text{معامل التحديد } (R^2) = \frac{SSR}{SST}$$

$$\text{وقيمة معامل التحديد في مثالنا السابق } (R^2) = \frac{SSR}{SST} = \frac{2272}{2442} = 0.93$$

ومن أجل توضيح أكثر لـ $(R^2)$، يمكن التفكير بـ SST كمقياس لمدى جودة $y$ كمخمن لحجم المبيعات السنوية . وبعد توفيق خط الانحدار المقدر ، نحتسب SSE كمقياس لجودة $\hat{y}$ كمخمن لحجم المبيعات السنوية. وعليه فإن SSR - الفرق بين SST و SSE - يقيس حقيقة ذلك الجزء من SST المفسر من خلال توفيق خط الانحدار المقدر . وعليه فإننا نفكر بـ R كما في الصيغة التالية :

$$R^2 = \frac{\text{مجموع مربعات الإختلاف المفسّرة بخط الانحدار}}{\text{مجموع مربعات الإختلافات قبل خط الانحدار}}$$

وعندما يعبر عنها كنسبة مئوية ، فان $R^2$ كنسبة مجموع المربعات المفسرة بخط الانحدار المقدر. وفي مثالنا الماضي نستنتج أن معادلة خط الانحدار تفسر ما نسبته 93% من مجموع مربعات الكلي . ونكون في أشد السعادة إذا ما حصلنا على مثل هذه القيمة لـ $R^2$ ، حيث أن الباحثين في الدراسات الاقتصادية والتجارية يشعرون بنتائج جيدة إذا حصلوا على $R^2$ مساوية لـ 0.60 وأكثر .

**تكمن مساوي معامل ارتباط بيرسون بأنه لا يطبق إلاّ على المتغيرات الرقمية أو الكمية ويفقد مفعوله بالنسبة للمتغيرات الوصفية.**

### رابعا - الارتباط :

هناك بعض الحالات التي يكون فيها صانع القرار غير معني بالمعادلة التي تربط بين متغيرين أو بتخمين أو التنبؤ بالمتغير التابع إذا عرفنا المتغير المستقل. فهناك طريقة أو أسلوب إحصائي في هذه الحالات يستخدم للكشف عن وجود علاقة بين المتغيرين، ولقياس ومعرفة اتجاه ومقدار العلاقة بينهما إن وجدت، وهذا الأسلوب يسمى بالارتباط البسيط. ومن أهم مقاييس الارتباط ما يشار بمعامل الارتباط (R). ويكون هذا المعامل محصوراً بين 1+ و 1- يشير + 1 إلى أن العلاقة إيجابية (طردية) وكاملة بين المتغيرين ، وعندها ستكون جميع النقاط الموجودة في الرسم الانتشاري واقعة على خط مستقيم بميل إيجابي . أما القيمة - 1فتشير إلى أن العلاقة بين X و Y كاملة وبشكل سلبي أو عكسي وتقع جميع النقاط على خط مستقيم ذات ميل سالب .

إذن قد يكون الارتباط بين المتغيرين طردياً (موجباً) بمعنى أن زيادة قيم أحد المتغيرين تصاحبها زيادة في قيم المتغير الآخر. وقد يكون الارتباط عكسياً (سالباً) إذا كانت القيم الصغيرة لأحد المتغيرين تصاحبها قيم كبيرة للمتغير الآخر . وعليه فان معامل الارتباط يعكس خاصتين هامتين هما : اتجاه العلاقة (طردي أو عكسي)، ودرجة هذه العلاقة وقوتها. حيث تحدد الإشارات الجبرية اتجاه الارتباط بينما تعكس القيمة المطلقة درجة الارتباط .

ومن الجدير بالذكر أن معامل الارتباط لا يتناول موضوع العلاقة الدالية، بمعنى أنه يقيم مقدار العلاقة بين المتغيرين واتجاههما دون التعرض إلى موضوع أي منهما متغير مستقل وأيهما متغير تابع.

سنستعرض فيما يلي أهم الصيغ المستخدمة لاستنباط معامل الارتباط :

**1- تحديد معامل الارتباط من خلال تحليل الانحدار :**

من خلال مناقشتنا السابقة للانحدار الخطي كنا قد افترضنا أن معادلة الانحدار الخطي بطريقة المربعات الصغرى هي $\hat{y} = a + bx$ . وفي هذه الحالة فان معامل الارتباط يمكن احتسابه من معامل التحديد $(r^2)$ كما يلي :

$$R = \pm\sqrt{r^2}$$

$$= \pm\sqrt{\text{معامل التحديد}}$$

وتتحدد إشارة معامل الارتباط من خلال إشارة الميل (b) في معادلة خط الانحدار . وفي مثالنا السابق فان معامل الارتباط هو

$$r = \pm\sqrt{0.93}$$

$$= +0.96$$

وحيث أن الميل (b = 4) كان موجباً فان معامل الارتباط هو إيجابي أيضاً.

**2- معامل الارتباط البسيط (بيرسون):**

هو معامل يمكن اللجوء إليه عندما لا يكون صانع القرار معنياً بالعلاقة الدالية بين المتغيرين x و y ، بل يكون معني فقط العلاقة بين المتغيرين من عدمها . فيمكن احتساب معامل الارتباط دون اللجوء إلى إنجاز تحليل خط الانحدار. وفي هذه الحالة فان الصيغة المستخدمة هي :

$$R = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x^2_i - (\sum x_i)^2}\sqrt{n\sum y^2_i - (\sum y_i)^2}}$$

وفي مثالنا السابق حول حجم المبيعات السنوية فان الحسابات الضرورية لاستخدام الصيغة أعلاه موضحة في الجدول (7) التالي :

جدول (7): الحسابات الضرورية لاستخراج معامل الارتباط

| مندوب المبيعات | سنوات الخبرة $x_i$ | حجم المبيعات $y_i$ | $x_i y_i$ | $x^2_i$ | $y^2_i$ |
|---|---|---|---|---|---|
| 1 | 1 | 80 | 80 | 1 | 6400 |
| 2 | 3 | 97 | 291 | 9 | 9409 |
| 3 | 4 | 92 | 368 | 16 | 8464 |
| 4 | 4 | 102 | 408 | 16 | 10404 |
| 5 | 6 | 103 | 618 | 36 | 10609 |
| 6 | 8 | 111 | 888 | 64 | 12321 |
| 7 | 10 | 119 | 1190 | 100 | 14161 |
| 8 | 10 | 123 | 1230 | 100 | 15129 |
| 9 | 11 | 117 | 1287 | 121 | 13689 |
| 10 | 13 | 136 | 1768 | 169 | 18496 |
| المجموع | 70 | 1080 | 128,8 | 632 | 082,119 |

وبتطبيق الصيغة الرياضية أعلاه فان :

$$R = \frac{10(8,128) - 70(1,080)}{\sqrt{10(632) - (70)^2}\sqrt{10(119,082) - (1,080)^2}}$$

ويلاحظ أن قيمة معامل الارتباط التي حصلنا عليها هي نفس القيمة التي احتسبت كجذر تربيعي لمعامل التحديد .

وهناك خصائص مميزة يتمتع بها معامل بيرسون ، تسهل من عملية احتسابه وهي :

- أنه لا يتأثر بإضافة أو طرح أي ثابت على/من جميع قيم أي من المتغيرين .
- كذلك فانه لا يتأثر بضرب جميع قيم أي من المتغيرين في أية كمية ثابتة .

ولتوضيح أهمية هاتين الخاصيتين سنحاول طرح مثال مبسط يمكن من خلاله استغلال وتسهيل عملية الاحتساب للصيغة المذكورة لارتباط بيروسون .

لمعرفة الارتباط بين حجم المبيعات ونفقات الدعاية والإعلان في أحد المؤسسات، أخذت عينة لفترة محدودة خمسة شهور حيث كانت فيها نفقات الدعاية وحجم المبيعات كما يلي في الجدول التالي :

جدول (8): الحسابات الضرورية لاستخراج معامل الارتباط

| الشهر | نفقات للدعاية $x_i$ | حجم المبيعات بالألف $y_i$ | $x_i y_i$ | $y^2_i$ | $x^2_i$ |
|---|---|---|---|---|---|
| 1 | 2 | 60 | 120 | 3600 | 4 |
| 2 | 5 | 100 | 500 | 10000 | 25 |
| 3 | 4 | 70 | 280 | 4900 | 16 |
| 4 | 6 | 90 | 540 | 8100 | 36 |
| 5 | 3 | 80 | 240 | 6400 | 9 |
| المجموع | 20 | 400 | 1680 | 33000 | 90 |

وبتطبيق الصيغة المذكورة لمعامل بيرسون

$$R = \frac{5(1680) - 20(400)}{\sqrt{5(90) - (20)^2}\ \sqrt{5(33300) - (400)^2}}$$

$$= \frac{8,400 - 8,000}{\sqrt{50 \text{ X } 5,000}} = \frac{400}{500}$$

$$= 0.8$$

هذا يعني أن العلاقة بين نفقات الدعاية وحجم المبيعات طردية ، فكلما زادت نفقات الدعاية كلما زادت المبيعات .

وإذا ما استخدمنا أحد الخاصيتين المذكورتين بقسمة (Y) على 10 فإن النتائج ستكون كالتالي :

جدول (9): الحسابات الضرورية لاحتساب معامل الارتباط بعد قسمة أحد المتغيرين (y) على 10

| الشهر | $X_i$ | $Y_i$ | $Y_i/10$ | $X^2_i$ | $(Y_i/10)^2$ | $X_i(Y_i/10)$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 60 | 6 | 4 | 36 | 12 |
| 2 | 5 | 100 | 10 | 25 | 100 | 50 |
| 3 | 4 | 70 | 7 | 16 | 49 | 28 |
| 4 | 6 | 90 | 9 | 36 | 81 | 54 |
| 5 | 3 | 80 | 8 | 9 | 64 | 24 |
| المجموع | 20 | 400 | 40 | 90 | 330 | 168 |

وبتطبيق المعادلة المذكورة أعلاه حصلنا على معامل الارتباط = +0.8 وهي نفس النتيجة التي حصلنا عليها عند استخدام العمود (Y) الأصلي .

### 3- معامل ارتباط سبيرمان (ارتباط الرتب) Rank Correlation

يقدم معامل بيرسون مقياساً ممتازاً للارتباط ولكنه يعتبر أفضل ما يكون في ظروف معينة ويفقد ميزته تحت ظروف أخرى حيث لا يصلح للاستخدام إذا كان المتغير لا يمكن قياسه مثل تقديرات الطلاب حيث يمكن أن تعرف أن (أ) أفضل من (ب) دون معرفة بكم تماماً، فعندما تكون المعلومات في صورة كيفية أن تظهر أو تصف وضعاً معيناً كالحالة الزواجية أو التعليمية أو حالة الطقس مثلاً، وفي هذه الحالات يتم اللجوء إلى بديل أكثر كفاءة يعتمد على ترتيب قيم كل متغير بينها وبين نفسها ثم تطرح الرتب من بعضها لتعرف ب

جدول (2): حساب مجموع مربعات الفروق (الاختلافات) لخط الانحدار المقدر $\hat{y} = 78 + 4.2\,x$

| الشهر | نفقات الدعاية $X_i$ | ترتيب قيم $X_i$ | حجم المبيعات $Y_i$ | ترتيب قيم $X_i$ | الفرق (d) | $D^2$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 5 | 60 | 5 | 0 | 0 |
| 2 | 5 | 2 | 100 | 1 | 1 | 1 |
| 3 | 4 | 3 | 70 | 4 | 1 | 1 |
| 4 | 6 | 1 | 90 | 2 | 1 | 1 |
| 5 | 3 | 4 | 80 | 3 | 1 | 1 |
| المجموع | | | | | | 4 |

(d = فروق الرتب) ويحسب معامل ارتباط الرتب حسب المعادلة التالية :

$$r = 1 - \frac{6\sum d^2}{n\,(n^2 - 1)}$$

بتطبيق هذه المعادلة على المثال السابق (العلاقة بين الإنفاق على الدعاية وبين المبيعات) .

$$r = 1 - \frac{6 \times 4}{125 - 5}$$

$$= 1 - \frac{24}{120}$$

$$= 1 - 0.2$$

$$= 0.8$$

وعادة ما يعطي بيرسون وسبيرمان قيماً متقاربة . ولو أن التطابق في هذا المثال لا يعدو أن يكون مجرد مصادفة . ويمكن التحقق بسهولة من أن معامل سبيرمان يتمتع كما يتمتع بيرسون بالخواص التي تم ذكرها لهذا المعامل وهي أنه لا يتأثر بطرح أو إضافة أو حتى قسمة أي من المتغيرين على أي ثابت.

مثال : عند احتساب العلاقة بين عدد الأطباء ومعدل الوفيات في دولة الكويت حسب طريقة سبيرمان كانت النتائج كما هي مبينة أدناه :

| السنوات | عددالأطباء $X_i$ | معدل الوفيات لكل ألف من السكان الكويتيين $Y_i$ | ترتيب $X_i$ | ترتيب $Y_i$ | الفرق (d) | $D^2$ |
|---|---|---|---|---|---|---|
| 1973 | 990 | 6.7 | 16 | 1 | 15 | 225 |
| 1974 | 1087 | 6.6 | 15 | 2 | 13 | 169 |
| 1975 | 1178 | 6.1 | 14 | 4 | 10 | 100 |
| 1976 | 1341 | 5.6 | 13 | 5 | 8 | 64 |
| 1977 | 1567 | 6.2 | 12 | 3 | 9 | 81 |
| 1978 | 1703 | 5.4 | 11 | 6.5 | 4.5 | 20.25 |
| 1979 | 1957 | 5.4 | 10 | 6.5 | 3.5 | 12.25 |
| 1989 | 2341 | 5.1 | 9 | 8 | 1 | 1 |
| 1981 | 2580 | 4.5 | 8 | 10 | 2 | 4 |
| 1982 | 2734 | 4.6 | 7 | 9 | 2 | 4 |
| 1983 | 2834 | 4.2 | 6 | 11 | 5 | 25 |
| 1984 | 2983 | 3.9 | 5 | 13 | 8 | 64 |
| 1985 | 3095 | 4.0 | 4 | 12 | 8 | 64 |
| 1986 | 3129 | 3.5 | 2 | 14 | 12 | 144 |
| 1987 | 3128 | 3.4 | 3 | 15.5 | 12.5 | 156.25 |
| 1988 | 3256 | 3.4 | 1 | 15.5 | 14.5 | 210.25 |
| المجموع | | | | | | 1344 |

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 1344}{(16)^3 - 16} = 1 - \frac{8064}{4096 - 16}$$

$$= 1 - \frac{8064}{4080} = -0.9764$$

وتدل هذه النتيجة على أن العلاقة قوية وعكسية بين عدد الأطباء ومعدل الوفيات ، فكلما زاد عدد الأطباء كلما انخفضت نسبة الوفيات.

### مثال تطبيقي على الحاسب الآلي:

كما هو معروف، فإن النظرية الاقتصادية ترى وجود علاقة طردية بين الانفاق الاستهلاكي و بين الدخل الشخصي المتاح. ويمكن صياغة ما ورد في النظرية الاقتصادية بالنموذج :

C=a+by يبين الجدول التالي بيانات الاستهلاك الخاص (Consumtion) والدخل المتاح (Income) في دولة ما في الفترة 1990 - 1981 (بالوحدات النقدية لتلك الدولة).

المطلوب إيجاد معاملات الانحدار الخطي البسيط، وهل هناك من علاقة جوهرية بين الاستهلاك والدخل، أي هل يمكن الاعتماد على العلاقة الخطية بالتنبؤ بقيمة الاستهلاك الشخصي إذا ما عرفنا قيمة الدخل (وذلك بمستوى 0.05) ؟.

| Year | Income | Cons |
|---|---|---|
| 1981 | 80 | 70 |
| 1982 | 100 | 65 |
| 1983 | 120 | 90 |
| 1984 | 140 | 95 |
| 1985 | 160 | 110 |
| 1986 | 180 | 115 |
| 1987 | 200 | 120 |
| 1988 | 220 | 140 |
| 1989 | 240 | 155 |
| 1990 | 260 | 150 |

SUMMARY OUTPUT

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.980847369 |
| R Square | 0.96206156 |
| Adjusted R Square | 0.957319256 |
| Standard Error | 6.493003227 |
| Observations | 10 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 1 | 8552.727273 | 8552.727 | 202.8679 | 5.75E-07 |
| Residual | 8 | 337.2727273 | 42.15909 | | |
| Total | 9 | 8890 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* |
|---|---|---|---|---|---|
| Intercept | 24.45454545 | 6.413817299 | 3.812791 | 0.005142 | 9.664247 |
| Income | 0.509090909 | 0.035742806 | 14.24317 | 5.75E-07 | 0.426668 |

من نتائج التحليل الوارد أعلاه فإن معاملات الانحدار هي:

a = 24.45

b = 0.51

أي أن معادلة خط الانحدار المقدرة هي :

C = 24.45 + 0.51Y

تشير النتائج أيضاً إلى أن العلاقة بين الاستهلاك والدخل علاقة جوهرية، حيث أن معامل الارتباط (R=0.98) كما أن معامل التحديد (R2=0.96)، ويعني ذلك أنه يمكن الاعتماد على العلاقة الخطية الواردة أعلاه في تقدير قيمة الاستهلاك عند معرفة قيمة الدخل المتاح.

## المصادر العربية


- د. علي أبو القاسم محمد، مقدمة في علم الإحصاء التطبيقي - المعهد العربي للتخطيط بالكويت.
- د. محمد عبدالحميد طه، مقدمة في الإحصاء - الهيئة العامة للتعليم التطبيقي والتدريب، 1985.
- د. علي أبو القاسم محمد، أساليب الإحصاء التطبيقي - المعهد العربي للتخطيط بالكويت، دار الشباب للنشر والترجمة، 1987.
- د. رمضان حسن عبدالرحيم، مبادىء في الإحصاء الوصفي، مكتبة عين شمس.
- د. أحمد عبادة سرحان ود. صلاح الدين طلبه، مقدمة الإحصاء الاجتماعي، دار الكتب الجامعية، 1973.
- جوردن بانكروفت وجورج أوسليفان، الرياضيات والإحصاء لدراسات المحاسبة والأعمال، ترجمة الدكتور سامي مقدسي، دار ماكجرو هيل للنشر، 1981.


## المصادر الانجليزية


- David Anderson, Dennis J. Sweeney and Thomas Williams, Introduction to Statistics, West Publishing Co., 1981.
- W.M. Harper, Statistics, Pitman Publishing, 1989.
- William L. Hays, Statistics for the Social Sciences, 1979.
- Ya-Lan Chau, Statistical Analysis for Business Economics, Elsevier Science Publishing Co., 1989.


المصادر العربية